Year: 2010

Proposal of the vote of thanks in discussion of Cule, M., Samworth, R., and Stewart, M.: Maximum likelihood estimation of a multidimensional logconcave density

Rufibach, K

Originally published at:
Rufibach, K (2010). Proposal of the vote of thanks in discussion of Cule, M., Samworth, R., and Stewart, M.: Maximum likelihood estimation of a multidimensional logconcave density. Journal of the Royal Statistical Society. Series B, 72(5):577-578.

Discussion of the paper

“Maximum likelihood estimation of a multidimensional logconcave density”

by Cule, Samworth & Stewart

Kaspar Rufibach
University of Zurich

May 2010

The authors are to be congratulated on the extension of log-concave density estimation to more than one dimension. Their work marks the temporary culmination of substantial research activity in shape-constrained density estimation over the last decade and directs the attention to (at least) two directions that previously had received little or no advertence. First, apart from very recent concurrent papers by Schuhmacher et al. (2009), Seregin and Wellner (2009), Koenker and Mizera (2010), Dümbgen et al. (2010), and Schuhmacher and Dümbgen (2010), nonparametric estimation of shape-constrained densities in dimension $d \geq 2$ has virtually received no attention. Apart from the theoretical obstacles related to these problems this neglect may be attributed to the difficulty of implementing algorithms to maximize the underlying likelihood function. The development of an algorithm and its implementation in `R` (Cule et al., 2009; R Development Core Team, 2009) for the log-concave case is certainly the first highlight of this paper. In dimension $d = 1$, after realizing that the maximizer of the likelihood function must be piecewise linear with kinks only at the observations, finding the log-concave density estimate boils down to maximizing a concave functional on $\mathbb{R}^n$ subject to linear constraints, see Rufibach (2007). In the multivariate case however, it is not clear how to parametrize the class of *concave* tent functions what hampers the formulation of a (linearly) constrained maximization problem similar to the univariate scenario. To circumvent this problem the authors modified the initial likelihood function to receive an updated functional whose *unconstrained* maximizer gives rise to the tent function that corresponds to the log-concave density estimate. This updated functional is concave but non-differentiable, disallowing the use of standard optimization algorithms. Instead, the authors successfully implemented (Cule et al., 2009) an algorithm due to Shor (Shor, 1985) which is able to handle non-differentiable target functionals.

As a second highlight the authors show that the estimator converges to the log-concave density f^* where this latter density minimizes the Kullback-Leibler divergence to f_0, the density of the observations. To the best of my knowledge, this general setup has not yet been considered for shape-constrained density estimation, not even for $d = 1$ e.g. in Groeneboom et al. (2001) or Dümbgen and Rufibach (2009). These latter papers only deal with the well-specified case where f_0 is log-concave. However, to assess robustness properties an analysis of the misspecified model is particularly valuable. A natural link here is: What are the limit distribution results under misspecification? In the well-specified univariate case, the pointwise limiting distribution is known, see Balabdaoui et al. (2009) and it seems worthwile to generalize these results to (1) higher dimensions and (2) the misspecified scenario.

Having shown consistency in some strong norms the natural next question is: What rates of convergence can be expected for the log-concave density estimator? For $d = 1$ rates of convergence, either in sup-norm (Dümbgen and Rufibach, 2009) or pointwise (Balabdaoui et al., 2009) have been derived.

If we assume that f_0 belongs to a Hölder-class with exponent $\alpha \in [1, 2]$, the minimax optimal rate for estimators within such a class can be derived from the entropy structure of the underlying function space and is $n^{-\alpha/(2\alpha+d)}$, cf. Birgé and Massart (1993). However, in the same paper it is shown that the rate of convergence for *minimum contrast estimators*, a class that contains maximum likelihood estimators (MLE) as a special case, is only $n^{-\alpha/(2d)}$ once $d > 2\alpha$.

The dependence of the exponents of these rates of convergence on dimension for $\beta = 2$, i.e. for densities with uniformly bounded second derivative, is displayed in Figure 1. The plot reveals that up to dimension $d = 4$ one can conjecture the MLE to be rate efficient, but beyond that split MLEs do not reach the minimax optimal rate anymore. Future work should aim at (1) in fact verifying the conjectured rates for the log-concave MLE in arbitrary dimension and (2) "fix" the log-concave MLE for dimensions $d > 4$ to make them also rate efficient for higher dimensions. How to achieve this goal is another open issue: (Additional) penalization comes to mind or consideration of classes of densities that are smaller than that of log-concave densities but yet nonparametric.

In addition to solving some important questions this paper has opened up new directions for research in shape-constrained density estimation and I am convinced that it will stimulate further research in the area. Consequently, I have great pleasure in proposing the vote of thanks.

References

Balabdaoui, F., Rufibach, K. and Wellner, J. A. (2009). Limit distribution theory for maximum likelihood estimation of a log-concave density. *Ann. Statist.* **37** 1299–1331.

Birgé, L. and Massart, P. (1993). Rates of convergence for minimum contrast estimators. *Probab. Theory Related Fields* **97** 113–150.
URL `http://dx.doi.org/10.1007/BF01199316`

Cule, M., Gramacy, R. and Samworth, R. (2009). LogConcDEAD: An R package for maximum likelihood estimation of a multivariate log-concave density. *Journal of Statistical Software* **29**.
URL `http://www.jstatsoft.org/v29/i02/`

Dümbgen, L. and Rufibach, K. (2009). Maximum likelihood estimation of a log-concave density and its distribution function. *Bernoulli* **15** 40–68.

Dümbgen, L., Samworth, R. and Schuhmacher, D. (2010). Approximation by log-concave distributions with applications to regression. Tech. rep., University of Bern.

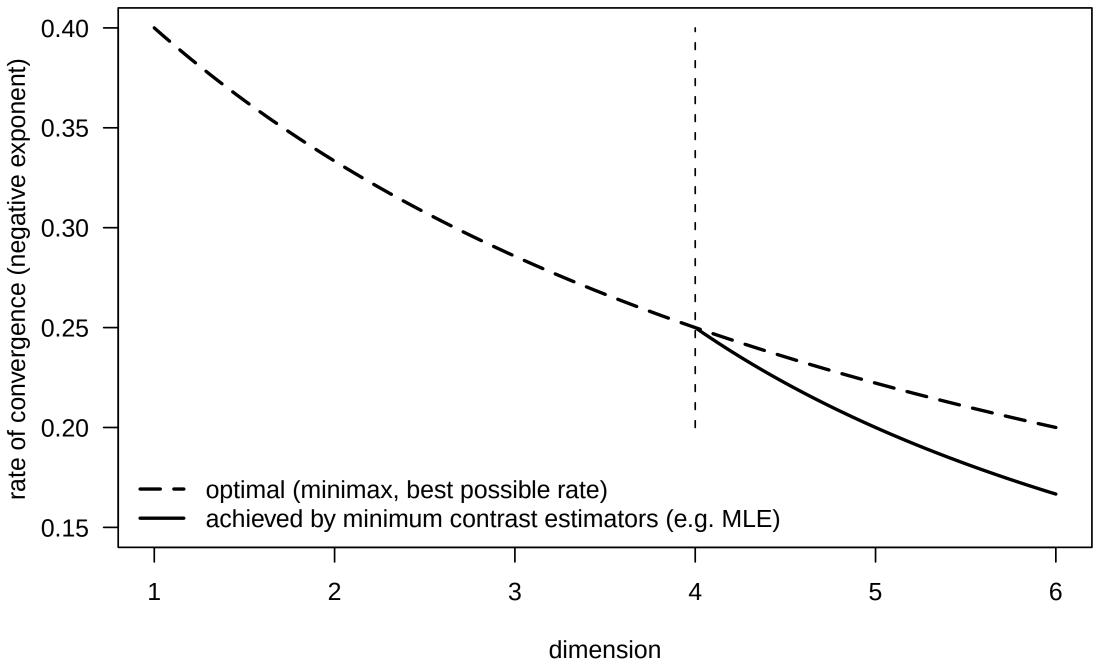

Figure 1: Rates of convergence for minimax optimal and minimum contrast estimators. Illustration for $\beta = 2$.

Groeneboom, P., Jongbloed, G. and Wellner, J. A. (2001). Estimation of a convex function: characterizations and asymptotic theory. *Ann. Statist.* **29** 1653–1698.

Koenker, R. and Mizera, I. (2010). Quasi-concave density estimation. *Ann. Statist.* to appear.

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
URL `http://www.R-project.org`

Rufibach, K. (2007). Computing maximum likelihood estimators of a log-concave density function. *J. Statist. Comp. Sim.* **77** 561–574.

Schuhmacher, D. and Dümbgen, L. (2010). Consistency of multivariate log-concave density estimators. *Statist. Probab. Lett.* **80** 376–380.

Schuhmacher, D., Huesler, A. and Duembgen, L. (2009). Multivariate log-concave distributions as a nearly parametric model. Tech. rep., University of Bern.
URL `http://www.citebase.org/abstract?id=oai:arXiv.org:0907.0250`

Seregin, A. and Wellner, J. A. (2009). Nonparametric estimation of multivariate convex-transformed densities. Tech. rep., University of Washington.
URL `http://www.citebase.org/abstract?id=oai:arXiv.org:0911.4151`

Shor, N. Z. (1985). *Minimization Methods for Non-Differentiable Functions*. Springer-Verlag, Berlin.